半滑舌鳎乙酰辅酶A羧化酶α基因全长cDNA分子克隆及饲料脂肪水平对其在肝脏中表达的影响

张夏青[1]　许建和[1,2]　潘　茜[3]　易乐飞[1]　彭永兴[1]　申　欣[1]　阎斌伦[1,2]　高　焕[1,2]　王雯祥[1]　程汉良[1,2*]

（1.淮海工学院，连云港 222005；2.江苏省海洋生物产业技术协同创新中心，连云港 222005；3.浙江省淡水水产研究所，湖州 313001）

摘　要：本试验采用RT-PCR（反转录PCR）和cDNA末端快速扩增（RACE）技术从半滑舌鳎肝脏中克隆了乙酰辅酶A羧化酶α（ACC1）基因全长cDNA，并采用实时荧光定量PCR方法对半滑舌鳎*ACC*1基因在肠道、肝脏、肌肉、卵巢、脾脏、全脑、肾脏、心脏、肠系膜脂肪等组织中的表达进行了研究；此外，还研究了饲料脂肪水平对半滑舌鳎肝脏中*ACC*1基因表达的影响。结果表明：1)半滑舌鳎*ACC*1基因cDNA全长7 811 bp，含1个7 074 bp的开放阅读框，编码2 357个氨基酸，ACC1蛋白计算分子质量为266 ku，等电点为6.42。半滑舌鳎ACC1氨基酸序列保守位点包括1个ATP结合位点(Gly^{316}～Gly^{321})、1个生物素结合位点(Val^{785}-Met^{786}-Lys^{787}-Met^{788})、1个辅酶A结合位点(Ser^{1969}～Val^{1995})。此外，半滑舌鳎*ACC*1基因存在可变剪接，形成另外2个同工型（isoforms），与分子质量为266 ku的ACC1相比，分别少8和15个氨基酸。半滑舌鳎ACC1氨基酸序列与其他脊椎动物的相似度很高，与斑马鱼和小鼠的相似度分别为93%和87%。2）半滑舌鳎所有组织中均检测到*ACC*1基因的表达，肝脏和全脑中*ACC*1 mRNA相对表达量显著高于其他组织（$P<0.05$），分别为2.67和2.53；肠道、卵巢和脂肪中次之，分别为1.14、0.97和1.10；肾脏中最低，仅为0.48。3）与对照组（未添加鱼油组）相比，3.5%鱼油组肝脏中*ACC*1 mRNA相对表达量显著降低（$P<0.05$）；7.0%和10.0%鱼油组肝脏中*ACC*1 mRNA相对表达量进一步降低，显著低于3.5%组和对照组（$P<0.05$），同时10.0%鱼油组低于7.0%鱼油组（$P>0.05$）。综上，本试验克隆出了半滑舌鳎*ACC*1基因的全长cDNA，并得出半滑舌鳎ACC1蛋白的主要功能位点为ATP结合位点、生物素结合位点、辅酶A结合位点，与其他脊椎动物相比基本保守。半滑舌鳎*ACC*1基因主要在肝脏和大脑等生脂组织中表达，饲料中添加鱼油显著抑制其肝脏中*ACC*1基因的表达，且抑制作用与鱼油添加量成正相关。

关键词：半滑舌鳎；乙酰CoA羧化酶α；分子克隆；营养调控；组织表达中图分类号：S963

文献标识码：A　　　　文章编号：

收稿日期：2015-08-01

基金项目：国家自然科学基金（31272636）；江苏省高校自然科学研究重大项目（10KJA240002）；江苏省自然科学基金（BK2012664）；江苏省海洋生物技术重点实验室开放基金（2009HS15）；浙江省重大科技专项（2012C12907-2）；江苏省优势学科建设工程项目（PAPD）；国家科技支撑计划（2012BAD26B04-04）

作者简介：张夏青（1990—），女，江苏宿迁人，硕士研究生，研究方向为水产动物营养与饲料。E-mail: zxq761114@163.com

*通信作者：程汉良，教授，硕士生导师，E-mail: CHL3139@163.com

乙酰辅酶 A 羧化酶（ACC，EC 6.4.1.2）催化乙酰辅酶 A 生成丙二酸单酰辅酶 A，是脂肪酸从头合成的限速酶[1-2]。原核生物 ACC 由 3 个基因编码的亚基组成，它们分别是生物素羧化酶（BC）、生物素羧基载体蛋白（BCCP）和羧基转移酶（CT）[3]。真核生物 ACC 由 1 个基因编码，并具有上述 3 个结构域。哺乳动物 ACC 有 α 和 β 2 种类型，它们具有不同的组织分布和功能。ACC-α 通常称为 ACACA 或 ACC1，主要在脂肪生成组织如肝脏中表达，催化长链脂肪酸的从头合成[4-6]。ACC-β 通常称为 ACACB 或 ACC2，是线粒体酶，主要在心脏和肌肉中表达，其产物丙二酰辅酶 A 通过抑制长链脂酰辅酶 A 从胞液向线粒体的转运而调控脂肪酸氧化[7-8]。在恒温动物中，已克隆出鼠[4]、鸡[5]、人[6]、牛[9]、绵羊[10]和山羊[11]*ACC*1 基因的全长 cDNA。人的 *ACC*1 基因含有 1 个 7 041 bp 的开放阅读框，编码 2 346 个氨基酸，ACC1 蛋白分子质量 255.5 ku；ACC2 蛋白分子质量为 280 ku。这 2 类 ACC 都由独立的基因进行编码，*ACC*1 基因在染色体 17q12 上，*ACC*2 基因在染色体 12q23 上，两者的总氨基酸序列的相似性达到 76%，两者的差别是在 N 端，ACC2 比 ACC1 多大约 140 个氨基酸，而 ACC2 就是通过这些多出的氨基酸序列锚定在线粒体的外膜上。关于鱼类 *ACC* 基因的研究较少，已克隆出草鱼 *ACC*1 基因全长 cDNA 序列，该基因主要在草鱼肝脏和脂肪组织中表达[12]。

半滑舌鳎是我国重要的海水养殖鱼类，但近年来其脂肪代谢相关疾病逐年增加，影响商品鱼肉质和成活率。本研究的目的是克隆半滑舌鳎 *ACC*1 基因全长 cDNA，分析其组织表达，并研究饲料中不同脂肪水平对其表达的影响，探讨该基因在脂肪代谢中的作用，为鱼类脂肪代谢营养调控研究提供基础资料和理论依据。

1　材料与方法

1.1　半滑舌鳎 *ACC*1 基因全长 cDNA 分子克隆

1.1.1　试验用鱼

试验用半滑舌鳎购自江苏省赣榆现代渔业示范园区，个体湿重 750~1 000 g，在淮海工学院海洋学院 240 L 水族箱中饲养 2 周，水温 20~23 ℃。

1.1.2　试验引物

本试验所用引物见表 1。各引物在 *ACC*1 cDNA 上的位置见图 1。

表 1　试验所用引物序列及预期产物大小

Table 1　Primer sequences and expected product size for the experiment

引物名称 Primer names	序列 Sequence（5′→3′）	用途 Use	预期产物大小 Expected product size/bp
反转录 PCR 引物 Reverse transcription PCR primers			
ACC-F_1	ATTGCSAACAATGGCATTGC	RT-PCR 扩增核心序列 1	592

ACC-R$_1$	CTTTTCCTCCTCCGCCCTC		
ACC-F$_2$	GAGGGCGGAGGAGGAAAAG	RT-PCR 扩增核心序列 2	1 246
ACC-R$_2$	CCTTCATGTAGGTGGTGTA		
ACC-F$_3$	GCCAGCAACATCACCTCTG	RT-PCR 扩增核心序列 3	727
ACC-R$_3$	GCTGGACGCTGTTSAGYTC		
ACC-F$_4$	TCAGTGACGGAGGTCTTCT	RT-PCR 扩增核心序列 4	732
ACC-R$_4$	TCAGCGTTGCAGCATGACT		
ACC-F$_5$	TGCCCAACTTCTTCTACCA	RT-PCR 扩增核心序列 5	2 530
ACC-R$_5$	GCCCAYGGCTGCATGATCTC		
ACC-F$_8$	TGCCCAACTTCTTCTACCA	*ACC*1 多态性识别	1 086～1 188
ACC-R$_8$	GGAGCGAACGGACTCTTC		
3′、5′cDNA 末端快速扩增引物 3′,5′RACE primers			
Oligo(dT)$_{16}$AP	CTGATCTAGAGGTACCGGATCC(T)$_{16}$	3′RACE 反转录合成第 1 链	
AP	CTGATCTAGAGGTACCGGATCC	与 ACC$_3$-F$_1$ 或 ACC$_5$-R$_2$ 合用	
Race$_3$-R	CTAGAGGTACCGGATCCTT	与 ACC$_3$-F$_2$ 合用	
ACC$_3$-F$_1$	AGGTTCTGGGATGATAGCA	与 AP 合用 3′RACE 第 1 轮	
ACC$_3$-F$_2$	CCCAGTGCCCATCCTTAT	与 Race$_3$-R 合用 3′RACE 巢式 PCR	1 520
ACC$_5$-R	GCTCCACATTAGCGTAGTTG	5′RACE 反转录合成第 1 链	
ACC$_5$-R$_1$	GCTCCACATTAGCGTAGTTG	与 Oligo(dT)$_{16}$AP 合用 5′RACE 第 1 轮	
ACC$_5$-R$_2$	ACGAATCGCCCTTTCATT	与 AP 合用 5′RACE 巢式 PCR	650
实时荧光定量 PCR 引物 RT-qPCR primers			
qACC-F	GGTGGAGTGTTGGAGCCTGAAG	*ACC*1 实时荧光定量 PCR	200
qACC-R	TGGTGGTAGATGGACAGCAGAAAC		
qRPL13A-F	CAATAAGGTTCTGCTGCTGGATGG	内参基因核糖体蛋白 L13A（ribosomal protein L13a，*RPL*13*A*）实时荧光定量 PCR	222
qRPL13A-R	CTGGGTGCTCTGAAGTGGTAAGG		
qRPL4-F	CATAGCGAGTGGAGCGTGGTTC	内参基因核糖体蛋白 L4（ribosomal protein L4，*RPL*4）实时荧光定量 PCR	203
qRPL4-R	GGGCTGGCGGCTGTTCTTG		
qPPIase-F	TCCACCGTCTCTGTTCTGTCTG	内参基因肽基脯氨酰顺反异构酶 A（peptidyl-prolyl cis-trans isomerase A，*PPIase*）实时荧光定量 PCR	176
qPPIase-R	CTTGTAGCCAAAGCCCTTCTCAC		
qB2M-F	TCCTCGTCTCTTCTGTGATCTTCC	内参基因 β2 微球蛋白（beta-2	140

qB2M-R	CCCAACCTTCTGTTCGCATCTAC	microglobulin，*B2M*）荧光定量 PCR

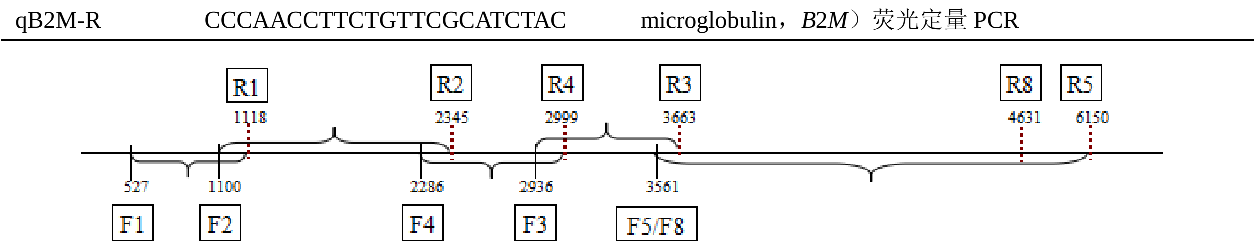


图 1　各引物在 *ACC*1 cDNA 上的位置

Fig.1　Locations of Primers in *ACC*1 cDNA

1.1.3　肝脏总 RNA 提取

半滑舌鳎活体解剖，快速分离肝脏，液氮研磨，使用 QIAGEN 公司的 RNeasy Lipid Tissue Mini Kit 试剂盒，按推荐方法提取肝脏总 RNA。1%的琼脂糖凝胶电泳检测 RNA 质量，核酸定量仪（Gene Quant）检测 RNA 的浓度后-70 ℃保存备用。

1.1.4　用于核心序列扩增的第 1 链 cDNA 反转录

采用 QIAGEN 的 QuantiTect Reverse Transcription Kit 反转录试剂盒，以 Oligo(dT)$_{16}$AP 为引物，按推荐方法去除 DNA 污染并反转录合成 cDNA 第 1 链，稀释 10 倍备用。

1.1.5　核心序列扩增

根据斑马鱼（XM_001919780）*ACC*1 基因序列，设计 5 对简并引物（表 1 和图 1），PCR 扩增 *ACC*1 基因 5 段核心序列，25 μL 反应体系含：TaKaRa 公司的 EmeraldAmp PCR Master Mix 12.5 μL，cDNA 第 1 链 1 μL，上、下游引物各 0.5 μL，ddH$_2$O 10.5 μL。扩增条件为：94 ℃变性 40 s、52 ℃退火 40 s、72 ℃延伸 60 s，共 35 个循环；反应前 95 ℃预变性 5 min；反应后 72 ℃充分延伸 7 min。1%的琼脂糖凝胶电泳检测，切胶回收，克隆测序。

1.1.6　半滑舌鳎 *ACC*1 基因 3′cDNA 末端快速扩增（RACE）

根据 *ACC*1 基因核心序列设计 3′RACE 特异引物（表 1）。以 cDNA 第 1 链为模板，ACC$_3$-F$_1$ 和 AP 为引物进行第 1 轮 PCR 扩增，将第 1 轮的 PCR 产物稀释 10 倍为模版，以 ACC$_3$-F$_2$ 和 Race$_3$-R 为引物进行第 2 轮扩增。3′RACE 的反应体系和反应条件同核心序列，1%的琼脂糖凝胶电泳检测，切取目的条带并回收，克隆测序。

1.1.7　半滑舌鳎 *ACC*1 基因 5′RACE

参考 Dieffenbach 等[13]的方法快速扩增 *ACC*1 基因 5′末端，5′RACE 扩增使用的引物见表 1。首先，以 ACC$_5$-R 为引物反转录合成 cDNA 第 1 链；然后，用 RNase H 分解 cDNA-RNA 杂交体中的 RNA，乙醇沉淀；再用末端脱氧核苷转移酶在 5′末端加 poly(A)尾，20 μL 反应体系如下：在有沉淀的原管中加入 5×Buffer 4 μL，dATP（100 mmol/L）1 μL，末端脱氧核苷转移酶（20 U/μL）1.5 μL，加 ddH$_2$O

至 20 μL。反应条件为：37 ℃温育 30 min，70 ℃灭活 10 min，用 TE 缓冲液稀释 10 倍后备用；最后，以 Oligo(dT)$_{16}$AP 和 ACC$_5$-R$_1$ 为引物进行第 1 轮 PCR 扩增，产物稀释 10 倍，用 Race$_3$-R 与 ACC$_5$-R$_2$ 进行第 2 轮 PCR 扩增，PCR 反应体系和条件同核心序列。1%琼脂糖凝胶电泳检测，切取目的条带并回收，克隆测序。

1.1.8 序列分析

用 DNAstar 7.1 软件包中的 SeqMan 对核心序列以及 3′和 5′末端序列进行组装，得到 *ACC*1 基因全长 cDNA 序列。用 EditSeq 对序列进行编辑和分析，寻找开放阅读框，并翻译成氨基酸序列。通过 http://www.cbs.dtu.dk/services/ 预测糖基化位点、磷酸化位点；通过 http://www.cbs.dtu.dk/services/SignalP/ 预测信号肽；通过 http://www.ebi.ac.uk/interpro/ 和 http://www.cbs.dtu.dk/services/TMHMM/预测跨膜结构域。使用 MEGA 6.0 采用邻接法(neighbor joining，NJ)法构建基于 ACC1 氨基酸序列的分子系统发育树[14]。

1.2 半滑舌鳎 *ACC*1 基因实时荧光定量 PCR

采用 SYBR Green 实时荧光定量 PCR 方法，对 *ACC*1 基因表达进行定量研究。

1.2.1 内参基因的克隆与定量引物设计

内参基因 β2 微球蛋白（beta-2 microglobulin，*B2M*）和核糖体蛋白 L13A（ribosomal protein L13a，*RPL*13*A*）从 GenBank 中获得，登录号分别为 FJ965562 和 GH232293; 核糖体蛋白 L4（ribosomal protein L4，*RPL*4）和肽基脯氨酰顺反异构酶 A（peptidyl-prolyl cis-trans isomerase A，*PPIase*）采用电子克隆方法获得，内参基因引物见表 1。

1.2.2 定量 PCR 检测体系

实时荧光定量 PCR 反应在 StepOnePlus PCR 仪（ABI）上进行。采用 QIAGEN 的 QuantiFast SYBR Green PCR Kit 定量试剂盒，20 μL 的反应体系中包含 2×PCR Master Mix 10 μL，25 mmol/L 的上、下游引物各 0.8 μL 和 10 倍稀释后的 cDNA 第 1 链 1 μL，ddH$_2$O 7.4 μL。采用 2 步 PCR 法进行扩增，首先 95 ℃预变性 5 min，然后是 40 个循环，每个循环包括 95 ℃变性 10 s 和 60 ℃延伸 30 s，循环结束后，从 60 ℃缓慢升温到 95 ℃，制备熔解曲线。每次反应都设置无模板对照，每个样品设 3 个技术重复。

1.2.3 半滑舌鳎 *ACC*1 基因的组织表达

取 3 尾喂食后 6 h 的半滑舌鳎活体解剖，快速分离肠道、肝脏、肌肉、卵巢、脾脏、全脑、肾脏、心脏、肠系膜脂肪等组织，分别提取总 RNA。采用 QIAGEN 的 QuantiTect Reverse Transcription Kit 反转录试剂盒，以 Primer Mix 为引物，按试剂盒推荐方法去除 DNA 污染并反转录合成 cDNA 第 1 链，稀释 10 倍用于荧光定量分析，经筛选确定 *RPL13A*、*RPL*4、*PPIase* 为组织表达的 3 个内参基因。

1.2.4 饲料脂肪水平对半滑舌鳎肝脏中 *ACC*1 基因表达的影响

1.2.4.1 试验饲料

共配制 4 种试验饲料，4 种试验饲料的鱼油添加量分别为 0（对照）、3.5%、7.0%和 10.0%（表 2）。制作饲料的鱼粉、鱼油、胆碱等主要原料由淮安通威饲料科技有限公司提供。根据配方，称量所需要的各种原料，充分混匀 10 min，加水后揉拌 15 min，再利用便携式绞肉机进行压缩制粒。将制粒的饲料均匀洒在白瓷盘内，放入干燥箱内 50 ℃烘 10 h，室温放置 1 h，装袋、编号、密封保存待用。

表 2 试验饲料组成及营养水平（干物质基础）

Table 2 Composition and nutrient levels of experimental diets (DM basis) %

项目 Items	饲料 Diets			
	1	2	3	4
原料 Ingredients				
进口红鱼粉 Imported brown fishmeal	60.0	60.0	60.0	60.0
酪蛋白 Casein	12.0	12.0	12.0	12.0
小麦淀粉 Wheat starch	24.0	20.5	17.0	14.0
磷酸二氢钙 $Ca(H_2PO_4)_2$	1.7	1.7	1.7	1.7
胆碱 Choline	0.3	0.3	0.3	0.3
鱼油 Fish oil		3.5	7.0	10.0
羧甲基纤维素 Carboxymethyl cellulose	1.0	1.0	1.0	1.0
预混料 Premix	1.0	1.0	1.0	1.0
合计 Total	100.0	100.0	100.0	100.0
营养水平 Nutrient levels[2)]				
粗蛋白质 Crude protein	50.0	50.0	50.0	50.0
粗脂肪 Crude fat	2.4	5.9	9.4	12.4

预混料为每千克饲料 The premix provided the following per kg of diets：VA 5 000 IU，VD 2 000 IU，VE 50 mg，VK 5 mg，VB_1 8 mg，VB_2 10 mg，VB_6 8 mg，VB_{12} 0.03 mg，泛酸 pantothenic acid 30 mg，烟酸 nicotinic acid 30 mg，生物素 biotin 0.4 mg，叶酸 folic acid 3 mg，VC 180 mg，肌醇 inositol 100 mg，Mg 300 mg，Fe 170 mg，Cu 4 mg，Zn 150 mg，Mn 22 mg，I 1 mg，Se 0.4 mg，Co 0.25 mg。

2）计算值 Calculated values。

1.2.4.2 饲养管理

试验在"单循环可控实验生态水槽系统"（大连汇新钛设备开发有限公司生产）内完成，每个水槽体积 240 L。整套系统有独立的水处理、温度控制、充气/增氧和紫外消毒等设备。试验用鱼购自江苏省赣榆现代渔业示范园区，共 240 尾，规格为 76.7~87.6 g/尾，随机分为 4 组，每组 3 个重复（水槽），

每个重复 20 尾，养殖周期 12 周，养殖水温 20~23 ℃。日投饲 2 次，投饲量随各水槽内半滑舌鳎的摄食情况及时调整，以 30 min 内吃完为度。

1.2.4.3　*ACC*1 基因相对表达量测定

养殖 12 周后，每组随机取 3 尾喂食后 6 h 的半滑舌鳎提取肝脏总 RNA，反转录合成 cDNA 第 1 链，采用 SYBR Green 实时荧光定量 PCR 方法测定 *ACC*1 基因的相对表达量。经筛选确定 *RPL*4 和 *B*2*M* 为内参基因。

1.2.5　数据分析

将内参基因的扩增效率和循环数（Cq）值数据导入 geNorm 程序，从而为不同的处理选择不同的内参组合。将目的基因和内参基因的扩增效率和 Cq 值数据导入 qBase Plus 程序，从而计算各基因在不同的处理下的相对表达量，结果采用平均值±标准差表示。统计分析采用 SPSS 13.0 软件，各组数据均通过了正态性和方差齐性检验，因此，组间差异采用单因素方差分析（one-way ANOVA），组间多重比较采用 Turkey’s 检验，P<0.05 表示差异显著。

2　结果与分析

2.1　半滑舌鳎 *ACC*1 基因全长 cDNA 分子特征

半滑舌鳎 *ACC*1 基因 cDNA 全长 7 811 bp（图 2），含 1 个 7 074 bp 的开放阅读框，编码 2 357 个氨基酸，ACC1 蛋白计算分子质量为 266 ku，等电点 6.42（GenBank 登录号为 KP033455)。此外，由于可变剪接，还发现了另外 2 个分子质量分别为 265 和 264 ku 的半滑舌鳎 ACC1 的同工型（isoforms）（GenBank 登录号分别为 KP033456 和 KP033457），与分子质量为 266 ku 的 ACC1 相比，分别少 8 和 15 个氨基酸。

```
     ---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+
1                                                                                                                           C     1
2    ATCATAATCATCACTCGTGGAGACGGGTATGTCTGTTAATCGGAGTTAACCTTTAACTGTCACAGGATTACGTGTCGGTTTATTCACGCCGTACTTGCTGCCTGTAAGTGTCACCCAGCC   121
122  ATGGCACAACAGGACAGCGGTGCCAAGAAGAATCCAGCAGTTGCAGCGCTGCACTCCCATTTCATCGTTGGCTCAGTGTCAGAGGAGAACTCCGAGGATGAAATCCAAGGGAAGCAGGAA   241
  1   M  A  Q  Q  D  S  G  A  K  K  N  P  A  V  A  A  L  H  S  H  F  I  V  G  S  V  S  E  D  N  S  E  D  E  I  Q  G  K  Q  E  40
     ←  Signal peptide
242  GTGCCGCTGGAGGAGAAGGAGACACGCTCTCTGTCGCCGTCATCTAGTAGCTCCGACAGCCCGTACGAAATAGTCTGTGACCAAAACGACGGCCCCACGCACAATCTAAGACCAAGTATG   361
 41   V  P  L  E  E  K  E  T  R  S  L  S  P  S  S  S  S  S  D  S  P  Y  E  I  V  C  D  Q  N  D  G  P  T  H  N  L  R  P [S  M] 80

362  TCTGGACTTCACCTAGTGAAGCAGGGCAGAGATCGCAGGCGTATGGACCTACAGAGAGACTTCACAGTTGCCTCACCAGCAGAGTTTGTCAAACGATTTGGTGGAAACAAGGTTATTGAG   481
 81  [S] G  L  H  L  V  K  Q  G  R  D  R  R  R  M  D  L  Q  R  D  F  T  V  A  S  P  A  E  F  V  K  R  F  G  G  N  K  V  I  E 120

482  AAGGTTCTCATTGCCAACAATGGCATTGCTGCTGTCAAATGTATGCGCTCCATACGACGATGGGCCTATGAGATGTTTCGGAATGAAAGGGCGATTCGTTTTGTTGTCATGGTGACGCCC   601
121   K  V  L  I  A  N  N  G  I  A  A  V  K  C  M  R  S  I  R  R  W  A  Y  E  M  F  R  N  E  R  A  I  R  F  V  V  M  V  T  P 160

602  GAAGACCTAAAAGCCAATGCAGAATACATCAAAATGGCAGATCACTATGTGCCTGTGCCGGGAGGGACCAACAACAACAACTACGCTAATGTGGAGCTCATTCTGGACATTGCTAAACGC   721
161   E  D  L  K  A  N  A  E  Y  I  K  M  A  D  H  Y  V  P  V  P  G  G  T  N  N  N  N  Y  A  N  V  E  L  I  L  D  I  A  K  R 200

722  ATACCTGTTCAGGCAGTGTGGGCTGGCTGGGGTCACGCCTCAGAGAACCCTAAACTTCCAGAGTTGCTGCATAAGAATGGCATTGCCTTCATGGGACCCCCAAGTCAGGCCATGTGGGCT   841
201   I  P  V  Q  A  V  W  A  G  W  G  H  A  S  E  N  P  K  L  P  E  L  L  H  K  N  G  I  A  F  M  G  P  P  S  Q  A  M  W  A 240

842  TTAGGAGACAAGATTGCCTCCTCCATTGTTGCTCAGACAGCTGGAATACCAACACTACCATGGAGTGGAGCAGGCCTGACAGTAGAATGGACAGAAAACGATCAAAAAAAGCAGACCATC   961
241   L  G  D  K  I  A  S  S  I  V  A  Q  T  A  G  I  P  T  L  P  W  S  G  A  G  L  T  V  E  W  T  E  N  D  Q  K  K  Q  T  I 280

962  TTAGGAGACAAGATTGCCTCCTCCATTGTTGCTCAGACAGCTGGAATACCAACACTACCATGGAGTGGAGCAGGCCTGACAGTAGAATGGACAGAAAACGATCAAAAAAAGCAGACCATC  1081
281   N  V  P  S  D  V  Y  E  R  G  C  I  K  D  V  E  D  G  L  T  A  S  E  K  V  G  Y  P  I  M  V  K  A  S  E [G  G  G  G  K]320

1082 GGAATCCGTAAAGTCAACACTGCTGATGATTTCCCTAACCTTTTCAGACAGGTCCAGACAGAGGTCCCGGGATCCCCGATTTTTGTCATGCAGCTTGCCAAACACGCTCGTCACTTGGAA  1201
321  [G] I  R  K  V  N  T  A  D  D  F  P  N  L  F  R  Q  V  Q  T  E  V  P  G  S  P  I  F  V  M  Q  L  A  K  H  A  R  H  L  E 360

1202 GTCCAGATCTTAGCTGATCAGTACGGCAACGCCATTTCCCTCTTTGGTAGAGATTGCTCCGTGCAACGACGACACCAAAAAATCATAGAGGAAGCTCCAGCTACTATCGCCACGTCTGAT  1321
361   V  Q  I  L  A  D  Q  Y  G  N  A  I  S  L  F  G  R  D  C  S  V  Q  R  R  H  Q  K  I  I  E  E  A  P  A  T  I  A  T  S  D 400

1322 GTGTTTGAGGATATGGAAAAGTGTGCAGTAAAGCTGGCTAAGATGGTGGGATACGTCAGCGCAGGTACAGTAGAATACTTGTACAGCCAGGATGGCAGCTTCTACTTCCTGGAGCTCAAC  1441
401   V  F  E  D  M  E  K  C  A  V  K  L  A  K  M  V  G  Y  V  S  A  G  T  V  E  Y  L  Y  S  Q  D  G  S  F  Y  F  L  E  L  N 440

1442 CCCCGTCTGCAGGTGGAACACCCCTGTACTGAGATGGTGGCCGACGTCAACTTGCCTGCCGCCCAATTGCAGATTGCAATGGGTATCCCCCTCTACCGGATCAAAGACATCAGGATGCTT  1561
441   P  R  L  Q  V  E  H  P  C  T  E  M  V  A  D  V  N  L  P  A  A  Q  L  Q  I  A  M  G  I  P  L  Y  R  I  K  D  I  R  M  L 480

1562 TACGGTGTCCAGCCTTGGGGAGACACTCACATCGACTTTGAGGGTCTGTCCACAGCGCCGTCACCACGGGGTCACGTCATCGCAGCACGAATCACAAGCGAAAATCCTGATGAGGGATTC  1681
481   Y  G  V  Q  P  W  G  D  T  H  I  D  F  E  G  L  S  T  A  P  S  P  R  G  H  V  I  A  A  R  I  T  S  E  N  P  D  E  G  F 520

1682 AAGCCAAGCTCTGGAACAGTGCAAGAGCTAAATTTCCGCAGCAACAAGAACGTGTGGGGCTACTTCAGTGTGGCGGCTGCTGGAGGGCTGCATGAGTTTGCTGACTCCCAGTTTGGACAC  1801
521   K  P  S  S  G  T  V  Q  E  L  N  F  R  S  N  K  N  V  W  G  Y  F  S  V  A  A  A  G  G  L  H  E  F  A  D  S  Q  F  G  H 560

1802 TGCTTCTCTTGGGGAGAGAATCGTGAAGAAGCCATCTCCAATATGGTGGTGGCTTTAAAGGAGCTGTCCATCAGAGGAGACTTCAGGACCACAGTGGAGTATCTCATAAAACTGCTAGAG  1921
561   C  F  S  W  G  E  N  R  E  E  A  I  S  N  M  V  V  A  L  K  E  L  S  I  R  G  D  F  R  T  T  V  E  Y  L  I  K  L  L  E 600

1922 ACAGAAAGCTTCCAACACAACAGCATTGACACAGGCTGGCTGGACAGACTTATCTCAGAGAAAATGCAGGCAGAGCGTCCTGATACCATGCTAGGAATCGTAAGTGGGGCGCTCCATGTG  2041
601   T  E  S  F  Q  H  N  S  I  D  T  G  W  L  D  R  L  I  S  E  K  M  Q  A  E  R  P  D  T  M  L  G  I  V  S  G  A  L  H  V 640

2042 GCGGATGTTAACTTCCGGAACAGTGTGTCTAATTTCCTGCATTCCTTGGAAAGGGGCCAAGTGCTGCCAGCACACACACTTCTCAATACCGTGGACGTGGAGCTAATCTACGAGGGCATA  2161
641   A  D  V  N  F  R  N  S  V  S  N  F  L  H  S  L  E  R  G  Q  V  L  P  A  H  T  L  L  N  T  V  D  V  E  L  I  Y  E  G  I 680

2162 AAGTATGTCCTAACGGTGACACGGCAGTCTCCAAACTCTTACGTGGTCATCATGAACGACTCATTGGCCGAGGTCGATGTCCATCGGCTCAGTGACGGAGGTCTTCTCCTGTCCTACGAT  2281
681   K  Y  V  L  T  V  T  R  Q  S  P  N  S  Y  V  V  I  M  N  D  S  L  A  E  V  D  V  H  R  L  S  D  G  G  L  L  L  S  Y  D 720

2282 GGCAGCAGCTACACTACGTACATGAAGGAGGAGGTGGACAGGTATCGCATCACTATTGGAAACAAGACGTGTGTATTTGAGAAGGAGAATGATCCGTCGCTGCTGCGATCTCCCTCAGCT  2401
721   G  S  S  Y  T  T  Y  M  K  E  E  V  D  R  Y  R  I  T  I  G  N  K  T  C  V  F  E  K  E  N  D  P  S  L  L  R  S  P  S  A 760

2402 GGAAAACTCATTCAGTTTACAGTGGAAGACGGCGGGCACCTATTCGCTGGCCAGTGCTATGCTGAAATAGAGGTGATGAAGATGGTAATGACTCTTACTACGGCAGAGTCTGGATGCATA  2521
761   G  K  L  I  Q  F  T  V  E  D  G  G  H  L  F  A  G  Q  C  Y  A  E  I  E [V  M  K  M] V  M  T  L  T  T  A  E  S  G  C  I 800

2522 CACTACGTCAAGAGGGCTGGAGCAGCACTGGAACCTGGCTGTGTCATTGCCAAGCTTCAGTTAGATGATCCGAGCAGAGTGCAGCAAGCAGAGCTGCACACTGGGGCCCTGCCGTCTCTC  2641
801   H  Y  V  K  R  A  G  A  A  L  E  P  G  C  V  I  A  K  L  Q  L  D  D  P  S  R  V  Q  Q  A  E  L  H  T  G  A  L  P  S  L 840

2642 CAGTCAGTTGCTTTGAGAGGGGAGAAGCTCCACAGAGTCTTCCACAACACACTTGACCACCTCGTCCACATCATGAATGGCTTCTGTCTTCCTGAATCGTTCTTCAACGCTAAGTTGAAA  2761
841   Q  S  V  A  L  R  G  E  K  L  H  R  V  F  H  N  T  L  D  H  L  V  H  I  M  N  G  F  C  L  P  E  S  F  F  N  A  K  L  K 880

2762 GAATGGGTGGAAAGGCTAATGAAGACCATGCGCGACCCCTCTTTACCTCTGTTGGAGCTACAAGACATCATGACCAGTGTGTCGGGTCGCATTCCCCCTGCCGTGGAGAAGTCCATCAAA  2881
881   E  W  V  E  R  L  M  K  T  M  R  D  P  S  L  P  L  L  E  L  Q  D  I  M  T  S  V  S  G  R  I  P  P  A  V  E  K  S  I  K 920

2882 AAGGAGATGGCCCAGTATGCCAGTAACATCACCTCTGTTCTCTGCCAGTTCCCAAGTCAGCAGATCGCCAACATCCTGGACAGTCATGCTGCAACGCTGAACAAGAAGTCTGAGAGGGAG  3001
921   K  E  M  A  Q  Y  A  S  N  I  T  S  V  L  C  Q  F  P  S  Q  Q  I  A  N  I  L  D  S  H  A  A  T  L  N  K  K  S  E  R  E 960

3002 GTGTTCTTCATGAACACACAGAGCATCGTGCAGCTGGTGCAGAAATATCGCAGTGGTATACGAGGTCACATGAAGGCAGTGGTGATGGACCTGCTCAGGCAGTACCTGAAAGTTGAGATC  3121
961   V  F  F  M  N  T  Q  S  I  V  Q  L  V  Q  K  Y  R  S  G  I  R  G  H  M  K  A  V  V  M  D  L  L  R  Q  Y  L  K  V  E  I 1000

3122 CAGTTCCAGAACGGGCATTACGACAAATGTGTGTTTGCGCTACGTGAGGAAAATAAAGTGGAGATGGCCAATGTACTCAACTATATCTTCTCCCATGCGCAAGTCACCAAGCGGAACTTG  3241
1001  Q  F  Q  N  G  H  Y  D  K  C  V  F  A  L  R  E  E  N  K  V  E  M  A  N  V  L  N  Y  I  F  S  H  A  Q  V  T  K  R  N  L 1040

3242 GTGGTCACGATGCTGATTGACCAGCTCTGTGGTCGTGAACAGACACTGACGGAGGAACTGGTGGCCATTTTGACTGAACTTACTCAGCTCAGCAAGACAACTAACGCCAAAGTAGCACTG  3361
1041  V  V  T  M  L  I  D  Q  L  C  G  R  E  Q  T  L  T  E  E  L  V  A  I  L  T  E  L  T  Q  L  S  K  T  T  N  A  K  V  A  L 1080

3362 CGTGCTCGACAGGTATTGATTGCTTCCCACCTTCCCTCCTACGAGCTGCGACACAACCAGGTGGAGTCCATCTTCCTGTCTGCTATAGATATGTACGGGCACCAGTTCTGCATTGAGAAC  3481
1081  R  A  R  Q  V  L  I  A  S  H  L  P  S  Y  E  L  R  H  N  Q  V  E  S  I  F  L  S  A  I  D  M  Y  G  H  Q  F  C  I  E  N 1120
                                                         ACC-F8   →
3482 CTGCAGAAACTGATCCTGTCAGAAACATCCATCTTCGATGTTTTGCCCAACTTCTTCTACCACAGCAATCAGGTAGTGAGGATGGCCGCACTGGAGGTCTACGTGCGCAGAGCGTACATT  3601
1121  L  Q  K  L  I  L  S  E  T  S  I  F  D  V  L  P  N  F  F  Y  H  S  N  Q  V  V  R  M  A  A  L  E  V  Y  V  R  R  A  Y  I 1160

3602 GCCTACGAGCTCAACAGCGTCCAGCACCGAGAGCTGAAGGACAACACGTGCATAGTGGAGTTCCAGTTCATGCTGCCGACCTCACATCCCAACAGAGGGAACATCCCCACACTTAACAGG  3721
1161  A  Y  E  L  N  S  V  Q  H  R  E  L  K  D  N  T  C  I  V  E  F  Q  F  M  L  P  T  S  H  P  N  R  G  N  I  P  T  L  N  R 1200

3722 ATGTCGTTCTCATCCAATCTGAACCACTACGGCATGGTACACGTAGCCAGTGTGAGCGATGTTCTGCTTGACACGTCGTTTACTCCGCCTTGTCAGCGAATGGGAGCGATGGTTGCGTTC  3841
1201  M  S  F  S  S  N  L  N  H  Y  G  M  V  H  V  A  S  V  S  D  V  L  L  D  T  S  F  T  P  P  C  Q  R  M  G  A  M  V  A  F 1240

3842 CGCTCCTTCCAGGACTTTGTCAGGAACATCACAGATGTGTTAAGTTGCTTCTCTGACTCTCCGCCCTCAAGTCCAACTTTCCCAGAGGGCGGCAATCCTGTTCTGTATGGTGAAGACGAT  3961
1241  R  S  F  Q  D  F  V  R  N  I  T  D  V  L  S  C  F  S  D  S  P  P  S  S  P  T  F  P  E  G  G  N  P  V  L  Y  G  E  D  D 1280

3962 AACAAGAGTGTCCAGGACGAACCGATCCACATCTTGAACGTAGCCATCAAGACAGACGGGGACATCGATGACGATAGCCTGGCAGCCATGTTCCGAGAGTTCACACTGTCAAAGAAATCC  4081
1281  N  K  S  V  Q  D  E  P  I  H  I  L  N  V  A  I  K  T  D  G  D  I  D  D  D  S  L  A  A  M  F  R  E  F  T  L  S  K  K  S 1320

4082 CTGCTGTTTGAACACGGAATACGAAGGCTGACGTTCCTCGTGGCTCAAAAGGATTTCAGGAAACAAATCAACTGTGAGATGGACCAAAGGTTTCACAGAGAATTCCCTAAATTCTTCACA  4201
1321  L  L  F  E  H  G  I  R  R  L  T  F  L  V  A  Q  K  D  F  R  K  Q  I  N  C  E  M  D  Q  R  F  H  R  E  F  P  K  F  F  T 1360
```

```
4202 TTCCGGGCGAGAGACAAATTTGAGGAGGACAGGATCTATCGTCATCTAGAGCCGGCGCTAGCTTTCCAGTTGGAGCTGAACCGCATGCGCAATTTCGCCCTAACCGCCATTCCGTGTTCC 4321
1361 F R A R D K F E E D R I Y R H L E P A L A F Q L E L N R M R N F A L T A I P C S 1400
4322 AACCACAAGATGCACCTTTACCTCGGTGCAGCCCGTGTGGAGGTGGGCACAGAGGTTACAGACTACCGCTTCTTTGTGCGAGCCATCATCCGCCACTCAGACTTGGTCACAAAGGAGGCT 4441
1401 N H K M H L Y L G A A R V E V G T E V T D Y R F F V R A I I R H S D L V T K E A 1440
4442 TCGTTTGAGTACCTCCACAATGAAGCCGAGCGGTTGCTGCTGGAAGCTATGGATGAGCTGGAGGTGGCTTTTAACAACACGACCGTCAGGACCGACTGTAACCACATCTTCCTGAACTTC 4561
1441 S F E Y L H N E A E R L L L E A M D E L E V A F N N T T V R T D C N H I F L N F 1480
                                         ← ACC-R8
4562 GTTCCTACAGTCATCATGGACCCGTCAAAGATCGAAGAGTCCGTTCGCTCCATGGTCATGCGCTACGGAAGCCGACTGTGGAAGCTGCGGGTTCTGCAGGCTGAGCTGAAAATCAACATC 4681
1481 V P T V I M D P S K I E E S V R S M V M R Y G S R L W K L R V L Q A E L K I N I 1520
4682 CGACTGACTCCGACAGGAAAGCAAATCCCCATCCGCCTGTTTCTGACCAACGAGTCTGGCTACTACCTCGACATCAGCCTTTACAAGGAGGTCACTGATTCCCGAACCGGACAGGTGGGG 4801
1521 R L T P T G K Q I P I R L F L T N E S G Y Y L D I S L Y K E V T D S R T G Q V G 1560
4802 CCCAAAGACCGACAGATCATGTTCCAGGCATATGGGGACAAACAGGGTCCACTTCATGGCATGCTCATAAACACACCCTACGTCACCAAGGACCTGCTGCAGTCCAAACGTTTCCAGGCA 4921
1561 P K D R Q I M F Q A Y G D K Q G P L H G M L I N T P Y V T K D L L Q S K R F Q A 1600
4922 CAGTCTCTAGGCACCACCTACGTCTATGACTTTCCTGAAATGTTCAGACAGGCTCTGAAAAAGTTGTGGCATAGTTGTCAGGCCTACGCTGAGTTACCCAAGTGTCCTCTACCCTCTGAG 5041
1601 Q S L G T T Y V Y D F P E M F R Q A L K K L W H S C Q A Y A E L P K C P L P S E 1640
5042 CTGCTCACCTTTACAGAGTTGGTCCTTGATGCCCAAGGTCAGCTGGTGCAGATGAACCGACTGCCGGGAGGCAACGAGATCGGAATGGTGGCTTGGCGGATGACCTTGCGGACACCCGAG 5161
1641 L L T F T E L V L D A Q G Q L V Q M N R L P G G N E I G M V A W R M T L R T P E 1680
5162 TGTCCGGCCGGACGTGAGATCGTCGTCATAAGTAATGATATCACGCACAAGATCGGATCGTTCGGGCCCCAGGAGGACTTGCTGTTCCTGCGAGCTTCGGAAATGGCCCGTGAGAGCGGC 5281
1681 C P A G R E I V V I S N D I T H K I G S F G P Q E D L L F L R A S E M A R E S G 1720
5282 ATCCCCCGGATCTACATCGCAGCCAACAGTGGCGCTCGCATCGGACTGGCAGAGGAAATAAGACACATGTTCCATGTGGCCTGGCATGATCCTTCTGACCCAAATAAGGGTTTCAAGTAC 5401
1721 I P R I Y I A A N S G A R I G L A E E I R H M F H V A W H D P S D P N K G F K Y 1760
5402 CTCTACCTCACACCTCAGGATTACAAAAAGGTTTCAGCCCTGAACTCTGTGCATTGTGAACACGTGGAGGACGAGGGAGAGTCCAGGTACAAGATCACTGACATTATTGGAAAGGATGAA 5521
1761 L Y L T P Q D Y K K V S A L N S V H C E H V E D E G E S R Y K I T D I I G K D E 1800
5522 GGACTCGGTGTGGAGAATCTCAAAGGTTCTGGGATGATAGCAGGAGAATCTTCTCTGGCTTATGACGAAATCATCACCATGAACCTGGTCACGTGTAGAGCCATAGGAATCGGGGCATAT 5641
1801 G L G V E N L K G S G M I A G E S S L A Y D E I I T M N L V T C R A I G I G A Y 1400
5642 TTGGTGAGGCTGGGACAAAGAACGATCCAGGTGGATAACTCTCACATTATTCTTACTGGAGCTGGAGCACTGAACAAGGTGCTGGGGAGAGAAGTATACACGTCAAACAATCAGCTGGGT 5761
1841 L V R L G Q R T I Q V D N S H I I L T G A G A L N K V L G R E V Y T S N N Q L G 1880
5762 GGAATACAGATCATGCACAACAACGGTGTGACTCACTGCACTGTCTGTGATGACTTTGAGGGAGTGTTCACCCTCCTGCAGTGGCTGTCCTACATGCCTAAGTGTACATCAAGCCCAGTG 5881
1881 G I Q I M H N N G V T H C T V C D D F E G V F T L L Q W L S Y M P K C T S S P V 1920
5882 CCCATCCTTATGGCCAAGGATCCCATTGATCGGCTCATAGACTTCATTCCCACCAAGGCCCCATATGACCCCCGCTGGATGCTGGCGGGGCGGCCCAGCCAGACTCCTAAGGGTTACTGG 6001
1921 P I L M A K D P I D R L I D F I P T K A P Y D P R W M L A G R P S Q T P K G Y W 1960
6002 CAGAGTGGTTTCTTTGATCAAGGATCTTTCATGGAGATTATGAAGCCGTGGGCCCAGAGTGTGGTGGTAGGCAGAGCCAGACTGGGCGGGATACCCACAGGAGTTGTTGCAGTGGAAACC 6121
1961 Q S G F F D Q G S F M E I M K P W A Q S V V V G R A R L G G I P T G V V A V E T 2000
6122 AGGTCAGTTGAGCTCTCAATACCAGCTGATCCTGCCAATCTGGACTCGGAGGCCAAAATCATCCAACAGGCAGGACAAGTGTGGTTCCCAGATTCTGCCTTCAAAACTGCACAAGCCATT 6241
2001 R S V E L S I P A D P A N L D S E A K I I Q Q A G Q V W F P D S A F K T A Q A I 2040
6242 AAGGACCTGAACCGAGGAGGCCTACCTCTCATAGTGTTCGCCAACTGGAGAGGGTTTTCTGGAGGCATGAAAGATATGTATGACCAGGTGTTGAAGTTTGGGGCCTATATCGTAGACGGG 6361
2041 K D L N R G G L P L I V F A N W R G F S G G M K D M Y D Q V L K F G A Y I V D G 2080
6362 CTGAGAGAATACAAACAGCCAGTGCTGGTTTACATTCCACCTCAGGCTGAACTAAGGGGAGGCTCTTGGGTGGTTATCGATCCTACCATCAACCCTCGACACATGGAGATGTACGCTGAC 6481
2081 L R E Y K Q P V L V Y I P P Q A E L R G G S W V V I D P T I N P R H M E M Y A D 2120
                              qACC-F →
6482 AAAGAAAGCCGTGGTGGAGTGTTGGAGCCTGAAGGAACAGTGGAGATAAAGTTCCGAAGGAAAGACCTGGTGAAGACCATGAGAAGAGTGGATCCGGTCTACATGGGCCTGGCTGAAAGA 6601
2121 K E S R G G V L E P E G T V E I K F R R K D L V K T M R R V D P V Y M G L A E R 2160
                                                                       ← qACC-R
6602 CTGGGAACCCCAGAGCTGAGCCCACCTGACCGCAAAGAGCTGGAGACCAAGTTGAAGGAACGGGAAGAGTTTCTGCTGTCCATCTACCACCAGGTGGCTGTGCAGTTTGCAGACCTCCAT 6721
2161 L G T P E L S P P D R K E L E T K L K E R E E F L L S I Y H Q V A V Q F A D L H 2200
6722 GACACCCCAGGGCGCATGCAGGAGAAGGGGGTAATCTCGGACATCCTTGAATGGCGAACATCCCGCCGGTTCTTCTACTGGCGTCTGCGGCGTCTGCTGCTGGAGGACACGGTGAAGAGG 6841
2201 D T P G R M Q E K G V I S D I L E W R T S R R F F Y W R L R R L L L E D T V K R 2240
6842 AAGATTCAAGCTGCTAACGGCGAACTGACGGACGGTCAGATCCAGGCCATGTTACGCCGCTGGTTCGTGGAGGCCGAGGGAGCTGTTAAGGCGTACCTCTGGGATAACAACGAAGAGGTG 6961
2241 K I Q A A N G E L T D G Q I Q A M L R R W F V E A E G A V K A Y L W D N N E E V 2280
6962 GTGGGCTGGCTGGAGAGACATCTAGCCGAAGAGGAGGGCGCCAGGTCTGTGATCGACGAGAACATCAAGTACATCAGGCGCGATCACATCCTCAAACAGATCCGCAGCCTGGTTCAAGCC 7081
2281 V G W L E R H L A E E E G A R S V I D E N I K Y I R R D H I L K Q I R S L V Q A 2320
7082 AACCCAGAGGTTGCCATGGATTCAATCGTGCACATGACCCAGCACATCTCACCCACACAGAGAACAGAGGTGGTTCGTATCCTGACTACTATGGAGACATCAGCAAACTCCTAGTTTGAA 7201
2321 N P E V A M D S I V H M T Q H I S P T Q R T E V V R I L T T M E T S A N S *     2357
7202 GCCAGCCATCTCCCTTCTCTCTCTCTTCCTTCCTTGCTGTCTGAGCCAGAGTTAAGGCCGCAGAACTTAGGCCGTGATGGAGGCTTTATCCATAATTCTACGTTTTCATTTGAGAACAGT 7321
7322 GATTTAGATTTTTTTTTAAAAAGCTTAGCTTCAGCATCTACACCCATGACAAATTTTAAATTTAGCATATGTTGACAGGGAGCTGATTAACAACGGTACAAAATGGATGTTGTATTCAG 7441
7442 GCGTATGATGAGGTCACGATAACTCGAACATGGTACATGACGAAGCTTCGATGGCGATCGCTGATGATATTTGCATTGTGGCCGACAGGATCAGGACACGCGTGTCTTAAACCAGCCAGA 7561
7562 CTAACATTTCTGTTAGTTGTTCTAAACCAGGCGTGGGCAAACTTTAGACTCGCGGGCCACAATTGTCCTAAAATTGACCAGGGGGGTGGGGGGGAGTGGTCAGGAGCAGATAGCTGGAGC 7681
7682 GTTCTAAGAACCAGGATACATGACATCTACAGGACATGCATAAAGGCTACAATGCTAAAAATAGCATTTGTATTTTTGGCGCCATCTATTGGGGAAATGTGGGTATTGCAGGAAAAAAAA 7801
7802 AAAAAAAAAA                                                                                                                 7811
```

图 2　半滑舌鳎 *ACC*1 基因全长 cDNA 核苷酸序列及其编码氨基酸序列

Fig.2　The nucleotide sequence and deduced amino acid sequences of *ACC*1 full-length cDNA in half-smooth tongue sole (*Cynoglossus semilaevis*)

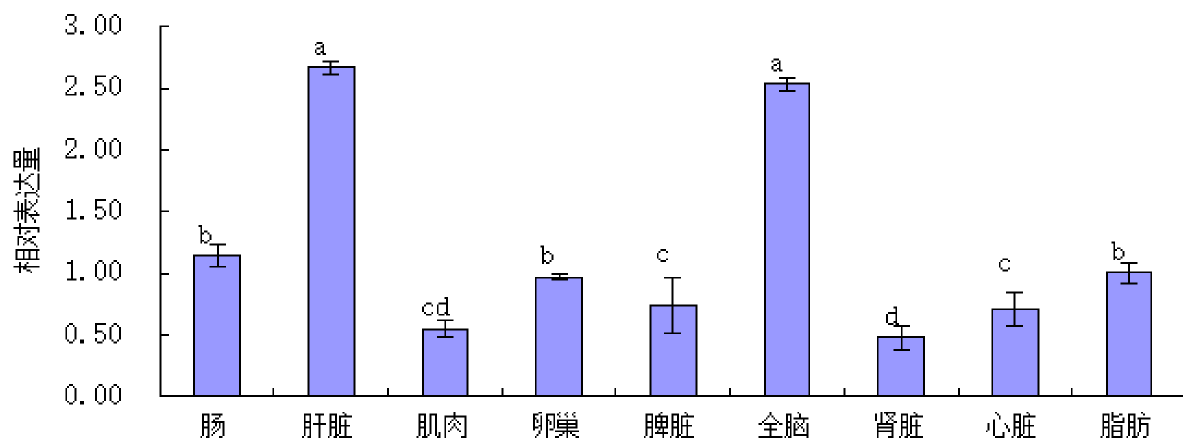
*代表终止密码子，灰色碱基代表荧光定量 PCR 和验证可变剪接的引物，灰色氨基酸表示信号肽，保守的磷酸化位点 Ser^{79}～Ser^{81}、潜在的 ATP 结合域 Gly^{316}～Gly^{321}、生物素结合位点 Val^{785}～Met^{788} 及 CoA 结合位点 Ser^{1969}～Val^{1995} 用氨基酸加框表示。

* represents a termination codon. The grey bases indicate the primers for RT-PCR and verifying alternative splicing. The

grey amion acids indicate signal peptides. Boxed amion acid sequences indicate the conservative phosphorylation sites Ser^{79} to Ser^{81}, the potential ATP-binding domain Gly^{316} to Gly^{321}, the biotin-binding tetrapeptide Val^{785} to Met^{788}, and the CoA-binding site Ser^{1969} to Val^{1995}.

2.2 半滑舌鳎 *ACC*1 基因的组织表达

以 *RPL13A*、*RPL*4、*PPIase* 3 个基因组合为内参基因，采用实时荧光定量 PCR 方法对半滑舌鳎 *ACC*1 mRNA 在肠道、肝脏、肌肉、卵巢、脾脏、全脑、肾脏、心脏、肠系膜脂肪等组织中的表达进行了研究。结果表明，所有组织中均检测到 *ACC*1 mRNA 的表达，肝脏和全脑中的相对表达量显著高于其他组织（$P<0.05$），其相对表达量分别为 2.67 和 2.53，肠道、卵巢和脂肪中相对表达量次之，分别为 1.14、0.97 和 1.10，肾脏中相对表达量最低，仅为 0.48（图 3）。



数据柱标注不同字母示表示差异显著（$P<0.05$）。图 5 同。

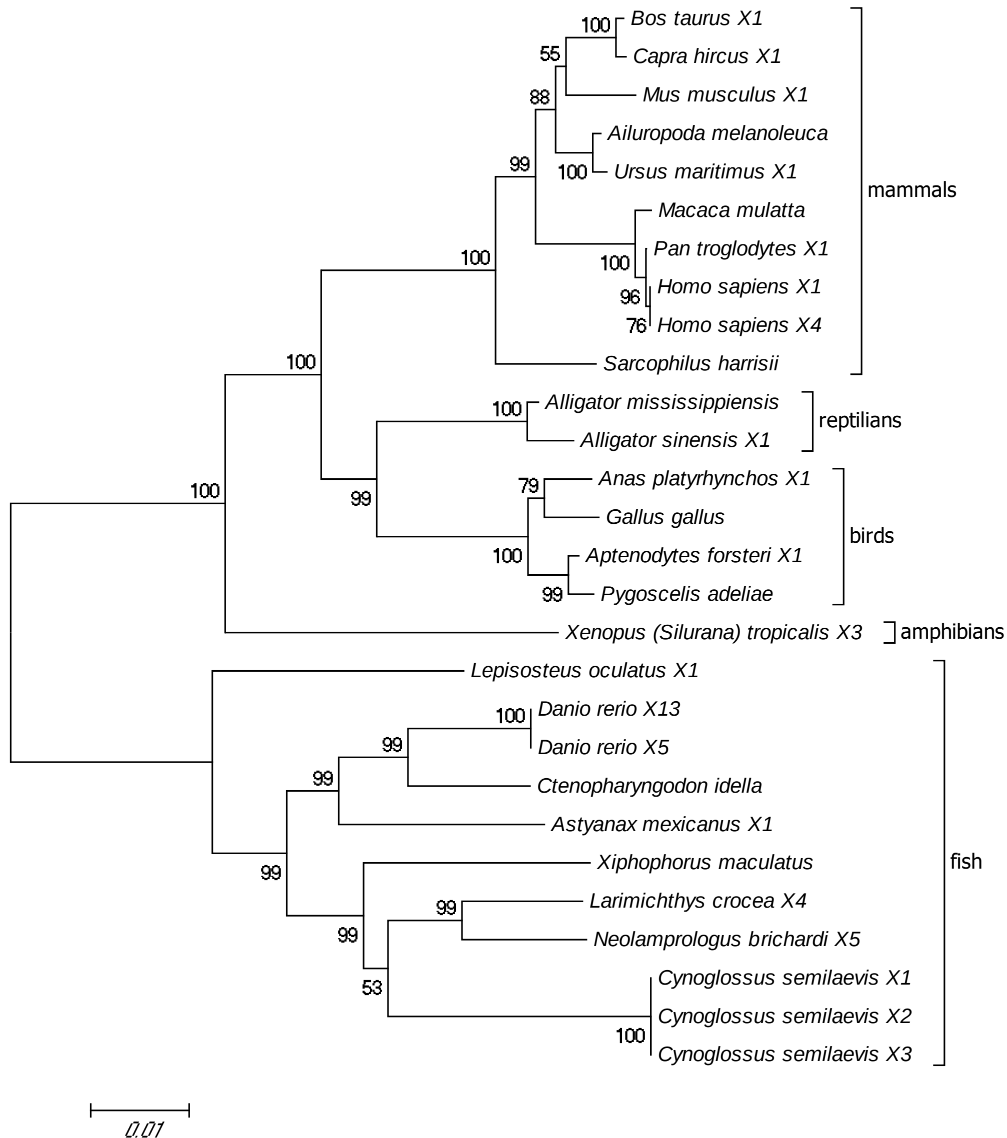
Value columns with different letters mean significant difference ($P<0.05$). The same as Fig.5.

图 3　半滑舌鳎 *ACC*1 基因组织表达分析

Fig.3　Expression analysis of *ACC*1 gene in different tissues of half-smooth tongue sole　（*Cynoglossus semilaevis*）

2.3 基于 ACC1 氨基酸序列的系统发育分析

从 GenBank 下载了 7 种鱼类和 16 种其他脊椎动物 ACC1 氨基酸序列用于系统发育分析。使用 MEGA 6.0 软件的 Clustal W 方法进行氨基酸序列比对，用 NJ 法构建基于 ACC1 氨基酸序列的系统发育树，用自展法（bootstrap method）进行系统发生检测，结果如图 4 所示。



节点上的数字是自展值，物种拉丁文后接的“X+阿拉伯数字”，表示该物种 ACC1 不同的同工型。The number on the node was the bootstrap value, and the “X+Arabic numeral” behind the Latin name mean the species’ different isoforms.

mammals：哺乳类；reptilians：爬行类；birds：鸟类；amphibians：两栖类；fish：鱼类；*Bos taurus*：牛；*Capra hircus*：山羊；*Mus musculus*：小鼠；*Ailuropoda melanoleuca*：大熊猫；*Ursus maritimus*：北极熊；*Macaca mulatta*：普通猕猴；*Pan trolodytes*：黑猩猩；*Homo sapiens*：智人；*Sarcophilus harrisii*：袋獾；*Alligator mississippiensis*：美国短吻鳄；*Alligator sinensis*：扬子鳄；*Anas platyrhynchos*：绿头鸭；*Gallus gallus*：原鸡；*Aptenodytes forsteri*：帝企鹅；*Pygoscelis adeliae*：阿德利企鹅；*Xenopus (Silurana) tropicalis*：非洲爪蟾；*Lepisosteus oculatus*：眼斑雀鳝；*Danio rerio*：斑马鱼；*Ctenopharyngodon idellus*：草鱼；*Astyanax mexicanus*：墨西哥脂鲤；*Xiphophorus maculatus*：月光鱼；*Larimichthys crocea*：大黄鱼；*Neolamprologus brichardi*：布氏新灿鲷。

图 4　基于 ACC1 氨基酸序列的系统发育树

Fig.4　Phylogenetic tree based on ACC1 amino acid sequences

由图 4 可见，所有鱼类聚为一枝，与其他脊椎动物分开。在鱼类中，同属鲤科的斑马鱼（*Danio rerio*）

（GenBank 登录号为 XP_009300128 和 XP_009300133）、草鱼（*Ctenopharyngodon idella*）（GenBank 登录号为 ADT82650）、墨西哥脂鲤（*Astyanax mexicanus*）（GenBank 登录号为 XP_007245267）聚为一枝，自展值为 99%；本研究克隆的半滑舌鳎的 3 个同工型首先聚为一枝，然后与同属鲈形总目（Percomorpha）的大黄鱼（*Laimichthys crocea*）（GenBank 登录号为 XP_010748451）、布氏新灿鲷（*Neolamprologus brichardi*）（GenBank 登录号为 XP_006799564）和月光鱼（*Xiphophorus maculatus*）（GenBank 登录号为 XP_005803670）聚为一枝，自展值为 99%，表明它们的亲缘关系较近。其他动物中，哺乳类、鸟类和爬行类均单独聚为一枝，自展值均为 100%，其中鸟类首先和爬行类聚为一枝，然后再与哺乳类聚为一枝，最后与两栖类聚为一枝。在哺乳类中，偶蹄动物牛（*Bos taurus*）（GenBank 登录号为 XP_005220033）和山羊（*Capra hircus*）（GenBank 登录号为 XP_005693213）聚为一枝；熊科动物大熊猫(*Ailuropoda melanoleuca*)(GenBank 登录号为 XP_002924940)和北极熊(*Ursus maritimus*)（GenBank 登录号为 XP_00869366）聚为一枝；灵长类动物智人（*Homo sapiens*）（GenBank 登录号为 NP_942131 和 XP_006721916）、黑猩猩（*Pan trolodytes*）（GenBank 登录号为 XP_511428）和普通猕猴（*Macaca mulatta*）（GenBank 登录号为 NP_001253707）聚为一枝，自展值均为 100%。基于 ACC1 氨基酸序列的系统发育关系与传统分类基本一致。此外，由于 ACC1 存在多种形式的可变剪接，每种生物的 ACC1 均有多个同工型，如斑马鱼的同工型多达 13 个，但每种生物的同工型都首先聚在一起，表明这些同工型是 1 个基因的不同剪接方式造成的。

2.4 饲料脂肪水平对半滑舌鳎肝脏中 *ACC1* 基因表达的影响

由图 5 可见，。与对照组相比，3.5%鱼油组肝脏中 *ACC1* mRNA 相对表达量显著降低（$P<0.05$）；7.0%和 10.0%鱼油组肝脏中 *ACC1* mRNA 相对表达量进一步降低，显著低于对照组和 3.5%组（$P<0.05$），同时 7.0%和 10%鱼油组间虽差异不显著（$P>0.05$），但在数值上 10%鱼油组低于 7%鱼油组。

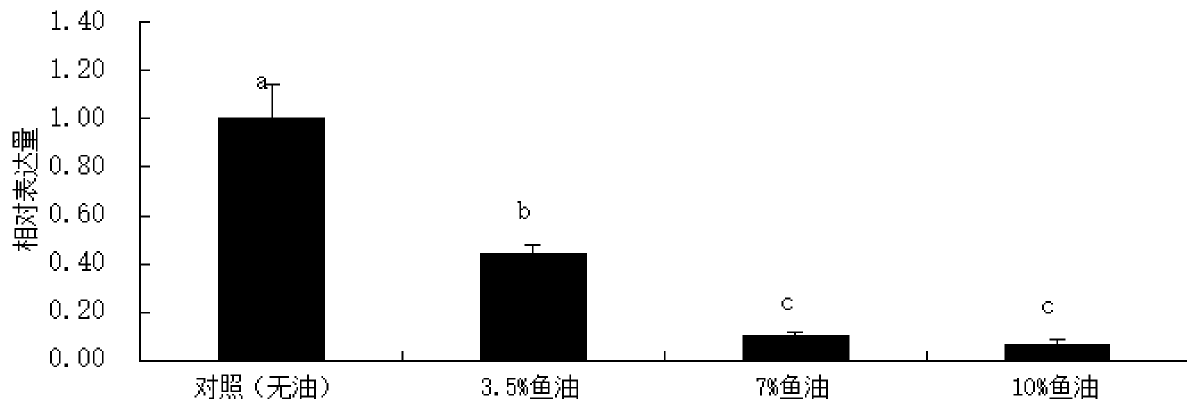


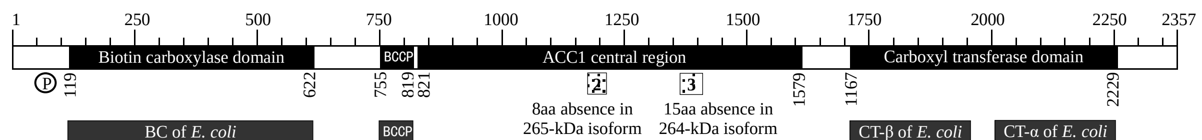
图 5 饲料脂肪水平对半滑舌鳎 *ACC*1 基因表达的影响

Fig.5 Effects of dietary fat level on *ACC*1 gene expression of half-smooth tongue sole （*Cynoglossus semilaevis*）

3 讨 论

3.1 半滑舌鳎 ACC1 蛋白结构

半滑舌鳎 ACC1 氨基酸序列与其他脊椎动物相比具有相似的结构特点，其主要结构域见图 6。



阴影框 2 和 3 表示可变剪接。Shaded boxes 2 and 3 showed alternative splicing.

Biotin carboxylase domain：生物素羧化酶结构域；ACC1 central region：ACC1 中心区域；Carboxyl transferase domain：羧基转移酶结构域；BCCP：生物素羧基载体蛋白，biotin carboxyl carrier protein；BC of *E. coil*：大肠杆菌 BC 结构域；CT-β of *E. coil*：大肠杆菌 CT-β 结构域；CT-α of *E.coil*：大肠杆菌 CT-α 结构域；8aa absence in 265-ku isoform：缺少 8 个氨基酸分子质量为 265 ku 的同工型：15aa absence in 264-ku isoform：缺少 15 个氨基酸分子质量为 264 ku 的同工型。

图 6 半滑舌鳎 ACC1 蛋白 3 个结构域与细菌 ACC1 蛋白的 3 个亚基比较

Fig.6 Comparison of three domains of ACC1 for half-smooth tongue sole and three subunits of ACC1 for bacteria

由图 6 可见，半滑舌鳎 ACC1 由 1 个基因编码，具有 3 个结构域，分别对应原核生物的 3 个基因——生物素羧化酶（BC）基因、生物素羧基载体蛋白（BCCP）基因和羧基转移酶（CT）基因，它们组成 *ACC*1 的 3 个亚基。BC 域包含 1 个 ATP 结合位点，甘氨酸富集区（GGGGKG）被认为是人 ACC1 和 ACC2 的 ATP 结合域[6-7]，半滑舌鳎的 ATP 结合位点是 Glu^{316}～Glu^{321}，与草鱼[12]和人类[6-7]ATP 结合域完全相同，说明这段序列高度保守。生物素酶的一个重要特征就是有 1 个生物素结合位点序列 Met-Lys-Met 的存在，这个区域有很保守的功能[3,15-16]。半滑舌鳎 ACC1 的 BCCP 域位于氨基酸序列的

第 755～819 位点，唯一的 4 肽序列 Val-Met-Lys-Met 位于氨基酸残基的第 785～788 位点（图 2）。因此，脊椎动物的生物素结合位点高度保守[17]。人 ACC1 中区域 SFSEIMQPWAQTVVVGRARLGGIPVGV 和 ACC2 中区域 SFKEIMAPWAQTVVTGRARLGGIPVGV 被认为是辅酶 A 结合位点[6-7]，在半滑舌鳎 CT 域中有相似的序列 ^{1969}SFMEIMKPWAQSVVVGRARLGGIPTGV^{1995}。

哺乳动物的 *ACC*1 基因共有 54 个编码蛋白质的外显子，在中心区域，编码 8 个氨基酸的外显子 E28 在大鼠[18]、小鼠[19]和绵羊[1]均存在可变剪接。本研究中，半滑舌鳎 ACC1 的中心区域位于氨基酸残基的第 821～1 579 位点（图 6），该区域存在可变剪接，形成分子质量分别为 266、265 和 264 ku 的 3 个同工型，与分子质量为 266 ku 的 ACC1 相比，分子质量为 265 和 264 ku 的 ACC1 分别少 8 和 15 个氨基酸，这与在大鼠[18]、小鼠[19]和绵羊[1]中发现的情况类似。

3.2　半滑舌鳎 *ACC*1 基因的组织表达

哺乳动物 *ACC*1 基因在各种组织中均表达，但在肝脏、脂肪、哺乳期乳腺等生脂组织中表达量最高[1]。半滑舌鳎 *ACC*1 在本试验检测的 10 种组织中均有表达，其中，肝脏和全脑等组织中表达量显著高于其他组织，与哺乳动物 *ACC*1 的组织表达规律相类似。

3.3　营养素对 *ACC*1 基因表达的调控

对高等动物的研究表明，饲料营养素可通过调控相关基因表达控制动物体脂的沉积，高糖饲料能促进脂肪的合成，其作用涉及基因转录、mRNA 加工和稳定多个层次[20]；脂肪酸作为转录因子的配体参与了动物体内多种脂肪代谢相关基因表达的调节[21]。研究表明，脂肪酸对生脂基因表达有抑制作用，但这种作用与脂肪酸种类关系密切，饲料中添加多不饱和脂肪酸(PUFA)可降低鼠肝脏 *ACC* 等生脂基因 mRNA 的相对表达量[22]。饲料中脂肪种类和水平对鱼类生脂基因表达有显著影响，黑鲷（*Sparus macrocephalus*）饲料中当 n-3 HUFA 含量在 0.92%以上时，脂肪酸合酶（*FAS*）基因表达粮显著下降[23]。草鱼饲喂分别以猪油和鱼油为主要脂肪源的饲料后发现，富含 n-3 HUFA 的鱼油组肝胰脏 *FAS*、*ACC* 等生脂基因 mRNA 的相对表达量显著低于猪油组[24]。鲫鱼肝胰脏 *ACC*1 和 *FAS* 等生脂基因 mRNA 相对表达量随饲料脂肪水平的升高而降低[25]。本研究中，饲料中添加鱼油显著抑制了半滑舌鳎肝脏 *ACC*1 基因的表达。因此，当饲料中提供充足的 PUFA 时，半滑舌鳎可以优先利用，从而减少脂肪酸的从头合成，以节省能量。

4　结　论

本研究从半滑舌鳎肝脏中克隆了 *ACC*1 基因全长 cDNA，半滑舌鳎 ACC1 的主要功能位点为 ATP 结合位点、生物素结合位点、辅酶 A 结合位点，与其他脊椎动物相比基本保守。半滑舌鳎 *ACC*1 基因主要在肝脏和大脑等生脂组织中表达，饲料中添加鱼油显著抑制其肝脏 *ACC*1 基因的表达，且抑制作用与鱼油添加量成正相关。

参考文献：

[1] BARBER M C,PRICE N T,TRAVERS M T.Structure and regulation of acetyl-CoA carboxylase genes of metazoa[J]. Biochimica et Biophysica Acta：Molecular and Cell Biology of Lipids,2005,1733(1):1–28.

[2] TONG L.Acetyl-coenzyme a carboxylase:crucial metabolic enzyme and attractive target for drug discovery[J].Cellular and Molecular Life Sciences,2005,62(16):1784–1803.

[3] JITRAPAKDEE S,WALLACE J C.The biotin enzyme family:conserved structural motifs and domain rearrangements[J].Current Protein & Peptide Science,2003,4(3):217–219.

[4] LÓPEZ-CASILLAS F,BAI D H,LUO X C,et al.Structure of the coding sequence and primary amino acid sequence of acetyl-coenzyme A carboxylase[J].Proceedings of the National Academy of Sciences of the United States of America,1988,85(16):5784–5788.

[5] TAKAI T,YOKOYAMA C,WADA K,et al.Primary structure of chicken liver acetyl-CoA carboxylase deduced from cDNA sequence[J].Journal of Biological Chemistry,1988,263(6):2651–2657.

[6] ABU-ELHEIGA L,ALMARZA-ORTEGA D B,BALDINI A,et al.Human acetyl-CoA carboxylase 2 Molecular cloning,characterization,chromosomal mapping,and evidence for two isoforms[J].Journal of Biological Chemistry,1997,272(16):10669–10677.

[7] ABU-ELHEIGA L,JAYAKUMAR A,BALDINI A,et al.Human acetyl-CoA carboxylase:characterization,molecular cloning,and evidence for two isoforms[J].Proceedings of the National Academy of Sciences of the United States of America,1995,92(9):4011–4015.

[8] ZHANG H L,YANG Z R,SHEN Y,et al.Crystal structure of the carboxyltransferase domain of acetyl-coenzyme A carboxylase[J].Science,2003,299(5615):2064–2067.

[9] MAO J,MARCOS S,DAVIS S K,et al.Genomic distribution of three promoters of the bovine gene encoding acetyl-CoA carboxylase alpha and evidence that the nutritionally regulated promoter I contains a repressive element different from that in rat[J].Biochemical Journal,2001,358(Pt.1):127–135.

[10] BARBER M C,TARVERS M T.Cloning and characterisation of multiple acetyl-CoA carboxylase transcripts in ovine adipose tissue[J].Gene,1995,154(2):271–275.

[11] BADAOUI B,SERRADILLA J M,TOMÀS A,et al.Goat acetyl-coenzyme a carboxylase α:molecular characterization,polymorphism,and association with milk traits[J].Journal of Dairy Science,2007,90(2):1039–1043.

[12] CHENG H L,JI N J,PENG Y X,et al.Molecular characterization and tissue-specific expression of the acetyl-CoA carboxylase α gene from grass carp,*Ctenopharyngodon idella*[J].Gene,2011,487(1):46–51.

[13] DIEFFENBACH C W,DVEKSLER G S.PCR primer:a laboratory manual[M].Plainview N Y:Cold Spring Harbor Laboratory Press,1995.

[14] TAMURA K,STECHER G,PETERSON D,et al.MEGA6:molecular evolutionary genetics Analysis version

6.0[J].Molecular Biology and Evolution,2013,30(12):2725–2729.

[15] JITRAPAKDEE S,ST MAURICE M,RAYMENT I,et al.Structure,mechanism and regulation of pyruvate carboxylase[J].Biochemical Journal,2008,413(3):369–387.

[16] LEE C K,CHEONG H K,RYU K S,et al.Biotinoyl domain of human acetyl-CoA carboxylase:structural insights into the carboxyl transfer mechanism[J].Proteins:Structure,Function,and Bioinformatics,2008,72(2):613–624.

[17] BAI D H,MOON T W,LOPEZ-CASILLAS F,et al.Analysis of the biotin-binding site on acetyl-CoA carboxylase from rat[J].European Journal of Biochemistry,1989,182(2):239–245.

[18] BARBER M C,POOLEY L,TRAVERS M T.Developmental regulation of alternatively spliced acetyl-CoA carboxylase-alpha mRNAs encoding isozymes with or without an eight amino acid domain upstream of the Ser-1200 phosphorylation motif in the mammary gland[J].Journal of Molecular Endocrinology,2001,27(3):349–356.

[19] SALLES J,SARGUEIL F,KNOLL-GELLIDA A,et al.Acetyl-CoA carboxylase and SREBP expression during peripheral nervous system myelination[J].Biochimica et Biophysica Acta：Molecular and Cell Biology of Lipids,2003,1631(3):229–238.

[20] KOO H Y,WALLING M A,CHUNG B H,et al.Dietary fructose induces a wide range of genes with distinct shift in carbohydrate and lipid metabolism in fed and fasted rat liver[J].Biochimica et Biophysica Acta：Molecular Basis of Disease,2008,1782(5):341–348.

[21] TAI C C,DING S T.N-3 polyunsaturated fatty acids regulate lipid metabolism through several inflammation mediators:mechanisms and implications for obesity prevention[J].Journal of Nutritional Biochemistry,2010,21(5):357–363.

[22] KAJIKAWA S,HARADA T,KAWASHIMA A,et al.Highly purified eicosapentaenoic acid prevents the progression of hepatic steatosis by repressing monounsaturated fatty acid synthesis in high-fat/high-sucrose diet-fed mice[J].Prostaglandins,Leukotrienes and Essential Fatty Acids,2009,80(4):229–238.

[23] 马晶晶.邵庆均,许梓荣，等.n-3 高不饱和脂肪酸对黑鲷幼鱼生长及脂肪代谢的影响[J].水产学报,2009,33(4):639-649.

[24] 李超.n-3 高不饱和脂肪酸对草鱼生长、脂代谢及健康状况的影响[D].硕士学位论文.杨凌：西北农林科技大学,2013.

[25] ZHOU J C,HAN D,JIN J Y,et al.Compared to fish oil alone,a corn and fish oil mixture decreases the lipid requirement of a freshwater fish species,*Carassius auratus gibelio*[J].Aquaculture,2014,428/429:272–279.

Full Length cDNA Molecular Cloning of Acetyl-CoA Carboxylase α Gene and Effects of Dietary Lipid Levels on Its Expression in the Livers of Half-Smooth Tongue Sole (*Cynoglossus Semilaevis*)

ZHANG Xiaqing[1] XU Jianhe[1,2] PAN Qian[3] YI Lefei[1] PENG Yongxing[1] SHEN Xin[1] YAN Binlun[1,2] GAO Huan[1,2] WANG Wenxiang[1] CHENG Hanliang[1,2*]

*Corresponding author, professor, E-mail: CHL3139 @163.com (责任编辑 菅景颖)

（1. *Huaihai Institute of Technology*, *Lianyungang* 222005, *China*; 2. *Co-Innovation Center of Jiangsu Marine Bio-Industry Technology*, *Lianyungang* 222005, *China*; 3. *Zhejiang Institute Of Freshwater Fisheries*, *Huzhou* 313001, *China*）

Abstract: The full-length cDNA of acetyl-coa carboxylase α (*ACC*1) gene was cloned from liver of half-smooth tongue sole (*Cynoglossus semilaevis*) by reverse transcription PCR (RT-PCR) and rapid amplification of cDNA ends (RACE) methods. The expression of *ACC*1 mRNA in gut, liver, muscle, ovary, spleen, brain, kidney, heart, mesenteric adipose tissue of half-smooth tongue sole was analyzed by real-time fluorescence quantitative PCR (RT-qPCR) method. In addition, the effects of dietary lipid level on *ACC*1 gene expression in liver of half-smooth tongue sole were investigated. The results showed as follows: 1) the full-length cDNA of *ACC*1 gene was 7 811bp with a 7 074 bp open reading frame encoding 2 357 amino acids. The ACC1 protein has a calculated molecular weight of 266 ku and isolectric point of 6.42. Some conserved sites of ACC1 amino acid sequence were found, including a ATP-binding site （Gly^{316} to Gly^{321}）, a biotin-binding site （Val^{785}-Met^{786}-Lys^{787}-Met^{788}）, and a CoA-binding site (Ser^{1969} to Val^{1995}). In addition, the alternative splice varieties were found in *ACC*1 gene, and we present evidence for the presence of two isoforms of *ACC*1 in half-smooth tongue sole liver that differ from the 266 ku ACC1 by the absence of 8 and 15 amino acids. The amino acid sequence has a high similarity to ACC1 of other species, the percent identity compared with zebrafish and mouse were 93% and 87%, respectively. 2) The expression of *ACC*1 mRNA was detected in all examined tissues. The relative expression level of *ACC*1 mRNA in the liver and brain were 2.67 and 2.53, respectively, which were significantly higher than that in other tissues (P<0.05); the relative expression level of *ACC*1 mRNA in gut, ovary and mesenteric adipose were second, which were 1.14, 0.97 and 1.10, respectively; the relative expression level of *ACC*1 mRNA in the kidney was the lowest, only was 0.48. 3) Compared with the control group (without fish oil group), diet added with 3.5% fish oil significantly decreased the relative expression level of *ACC*1 mRNA in liver (P<0.05); diet added with 7.0% and 10.0% fish oil led to further decrease the relative expression level of *ACC*1 mRNA in liver, and it significantly lower than that in 3.5% fish oil group and control (P<0.05), meanwhile, the 10.0% fish oil group was lower than 7.0% fish oil group (P>0.05). In summary, we have cloned the full-length cDNA of *ACC*1 gene from half-smooth tongue sole. Compared with other vertebrates, the main functional sites (ATP-binding site, biotin-binding site and CoA-binding site) are basically conserved. The lipogenic tissues, such as liver and brain, are the main ACC*1* gene expressing tissues in half-smooth tongue sole. Moreover, liver *ACC*1 mRNA expression is inhibited after the fish fed diets added with fish oil, and the inhibitory effect is positively correlated with the addition of fish oil.

Key words: half-smooth tongue sole (*cynoglossus semilaevis*); ACC1; molecular cloning; nutritional

regulation; tissue express